PREPRINT

PERTANIKA JOURNALS

TROPICAL AGRICULTURAL SCIENCE

Journal homepage: http://www.pertanika.upm.edu.my/

Genetic Variation and Phylogenetics of Strawberry Cultivars Grown in Indonesia Based on the *matK* Gene

Ganies Riza Aristya[1]*, Muhammad Fatih[1], Rina Sri Kasiamdari[2], Ani Widiastuti[3], and Muhammad Fauzi Arif[4]

[1]*Laboratory of Genetics and Breeding, Department of Tropical Biology, Faculty of Biology, Universitas Gadjah Mada, 55281 Yogyakarta, Indonesia*
[2]*Laboratory of Plant Systematics, Department of Tropical Biology, Faculty of Biology, Universitas Gadjah Mada, 55281 Yogyakarta, Indonesia*
[3]*Laboratory of Plant Pests & Disease, Department of Plant Pests and Diseases, Faculty of Agriculture, Universitas Gadjah Mada, 55281 Yogyakarta, Indonesia*
[4]*Laboratory of Biology, Department of Biology, Faculty of Mathematics and Natural Sciences, Mulawarman University, 75123 Samarinda, Indonesia*

ABSTRACT

Strawberries (*Fragaria* × *ananassa*) are a highly valued fruit crop widely cultivated in Indonesia. Molecular identification using the *matK* gene is widely applied for rapid and accurate genetic diversity analysis. The objectives of this study were to investigate genetic variation, genetic distance, haplotype distribution, and phylogenetic relationships among strawberry cultivars grown in Indonesia based on the *matK* gene, and to evaluate its usefulness for cultivar discrimination. Methods included sample collection, DNA extraction and quality assessment, DNA amplification of the *matK* gene applying specific *matK*3F and *matK*4R primers, sequencing, alignment with read-end trimming, detection of genetic variation and distance measurement, haplotype network construction, and phylogenetic tree construction using Maximum Likelihood (ML) and Bayesian Inference (BI). Fifteen strawberry cultivars were analyzed alongside ten reference sequences from databases. Results showed that all 15 cultivars had identical *matK* sequences in the 555-bp alignment and shared one haplotype with references of *F. virginiana*, *F. chiloensis*, *F. vesca* and *F. orientalis*. Interspecific distances were low (0.0036-0.0073), and the 25 ingroup sequences formed three haplotypes (Hd = 0.157). ML and BI trees were identical but weakly supported (bootstrap 62%; posterior probability 0.84). These results show that *matK* cannot distinguish these cultivars, consistent with their common maternal origin. As the first *matK* dataset for

ARTICLE INFO

Article history:
Received: 10 October 2025
Accepted: 18 November 2026
Published: 02 December 2026

DOI: https://doi.org/10.47836/pjtas.49.6.11

E-mail addresses:
ganies_riza@ugm.ac.id (Ganies Riza Aristya)
muhammad.fatih@mail.ugm.ac.id (Muhammad Fatih)
rkasiamdari@ugm.ac.id (Rina Sri Kasiamdari)
aniwidiastuti@ugm.ac.id (Ani Widiastuti)
fauzi.arif@fmipa.unmul.ac.id (Muhammad Fauzi Arif)
* Corresponding author

e-ISSN: 2231-8542

© Universiti Putra Malaysia Press

these cultivars, this study cautions against plastid barcodes for cultivar authentication. Further studies using nuclear markers are recommended to support breeding programs to develop superior local strawberry varieties.

Keywords: Chloroplast DNA, *Fragaria* × *ananassa*, genetic variation, haplotype, *matK* gene, phylogenetics

INTRODUCTION

Strawberry (*Fragaria* × *ananassa*) is a valuable fruit known for its sweet and sour flavor and high vitamin C content (Sapei & Hwa, 2014). The cultivated strawberry is a young allo-octoploid (2n = 8x = 56) that arose in eighteenth-century Europe from spontaneous hybridization between *F. virginiana* and *F. chiloensis* (Hardigan et al., 2021; Liston et al., 2014). Strawberries cultivated in Indonesia include introduced cultivars and local selections (Hanif & Ashari, 2012), grown mainly in highland areas such as West Bandung and Magelang (Arif et al., 2019; Azizah et al., 2019). National production reached 31,771 t in 2024, about 88% of which came from West Java (Direktorat Jenderal Hortikultura, 2025). Productivity is still limited by extreme weather and less-than-optimal cultivation practices (Siagian et al., 2016). As a result, there is an increasing need for plant breeding efforts intended to improve quality and productivity. This can be achieved through crossing and selection based on high genetic variation to produce superior varieties suited to local environments (Koryati et al., 2022; Vondracek et al., 2024).

Multiple molecular techniques have been used to assess the genetic diversity of strawberry cultivars grown in Indonesia. Aristya et al. (2019) used Random Amplified Polymorphic DNA (RAPD) markers, and Arif et al. (2019) used Cleaved Amplified Polymorphic Sequence (CAPS) markers to characterize cultivars from Magelang. The results showed that these methods could discover polymorphisms but were not optimal for comparing genetic differences in detail or across studies. As a result, DNA sequence-based markers are an attractive alternative, because the data are reproducible and can be compared directly with public databases. The plastid *matK* gene is one of the two core DNA barcodes for land plants (CBOL Plant Working Group, 2009) and is widely used to separate closely related species (Farooq et al., 2020; Savolainen et al., 2005; Udensi et al., 2017). In strawberry, Qarni et al. (2022) tested eight barcode regions in wild and cultivated *Fragaria* from Pakistan, but only ITS2 and *rbcL* were sequenced successfully.

Molecular analysis using the *matK* gene is widely used to measure genetic variation among plant species, as it can quickly and accurately differentiate between closely related species (Farooq et al., 2020; Savolainen et al., 2005; Udensi et al., 2017). This genetic variation allows the application of bioinformatics analyses such as genetic distance measurements to determine numerical differences between sequences (Joly et al., 2015), haplotype networks for visualizing intraspecific and biogeographic relationships (Leigh &

Bryant, 2015), and phylogenetic construction using the Maximum Likelihood and Bayesian Inference methods to visually understand the evolutionary relationships of *Fragaria* species in phylogenetic trees (Dharmayanti, 2011; San Mauro & Agorreta, 2010). However, plastid DNA is maternally inherited and shows low variation in *Fragaria* (Li et al., 2021; Njuguna et al., 2013). The octoploid strawberries, including *F.* × *ananassa*, share one maternal lineage (Njuguna et al., 2013; Song et al., 2023), although their nuclear genomes combine several diploid progenitors (Edger et al., 2019; Rousseau-Gueutin et al., 2009) and have been changed by breeding (Fan & Whitaker, 2024; Hardigan et al., 2021). Honjo et al. (2009) found only three maternal haplotypes among cultivars from North America, Europe and Japan. Whether *matK* can distinguish the cultivars grown in Indonesia has not been tested, and we expected little variation among them. This study aims to calculate genetic variation and genetic distance among strawberry cultivars grown in Indonesia, explore their haplotype distribution patterns and phylogenetic relationships based on the *matK* gene, evaluate its usefulness for cultivar discrimination, and give perspectives for breeding programs and biodiversity conservation. To our knowledge, this is the first study to report *matK* sequences for a set of strawberry cultivars grown in Indonesia and to test directly whether this barcode can separate them. The sequences serve as reference data for these cultivars, and the data offer practical guidance for choosing markers in strawberry breeding and germplasm work.

MATERIALS AND METHODS

Sample Collection and DNA Extraction

Fifteen strawberry cultivars were collected from Ija Mulyana Strawberry Garden (West Bandung Regency, West Java), Inggit Strawberry Garden 3 (Magelang Regency, Central Java) and UPTD BP2TPH (Sleman Regency, Special Region of Yogyakarta) (Figure 1). The origin of each cultivar is listed in Table 1. Young leaves were collected from three plants per cultivar and stored at -20 °C in the Laboratory of Genetics and Breeding, Universitas Gadjah Mada. In addition, ten *Fragaria matK* sequences and one outgroup sequence (*Potentilla indica*) were retrieved from GenBank (Table 1). Genomic DNA was isolated from leaves using the Genomic DNA Mini Kit Plant (Geneaid). DNA quality was tested by 0.8% agarose gel electrophoresis at 50 V for 45 min (Sambrook & Russell, 2001) and visualized using Gel Doc and Optilab. High-molecular-weight DNA (>10 kb) was considered intact (Arif et al., 2019; Sun et al., 2021). DNA concentration and purity (A260/A280) were measured with a NanoDrop spectrophotometer. A ratio of 1.8-2.0 indicates pure DNA (Sambrook & Russell, 2001); samples below this range were kept only if they produced a single, clear amplicon, which was used as the practical criterion of suitability for sequencing.

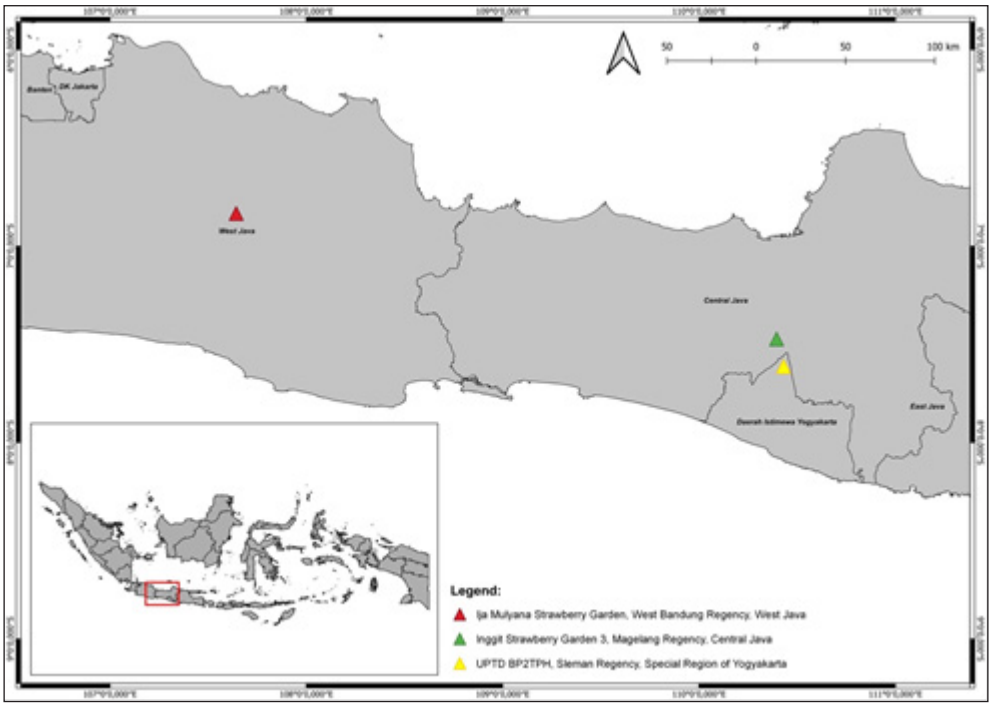


Figure 1. Sampling site locations in this study

Table 1
Strawberry cultivars used in this study, with their origin and GenBank accession numbers

Code	Cultivar or Species	Source of Collection	GenBank Accession
CL	*F.* × *ananassa* 'Californica'	Not documented	PV804952
KN	*F.* × *ananassa* 'Knia'	Not documented	PV804953
MR	*F.* × *ananassa* 'Mencir'	Indonesia (local cultivar)	PV804954
MH	*F.* × *ananassa* 'Moha'	Not documented	PV804955
GL	*F.* × *ananassa* 'Geolhyang'	Not documented	PV804956
JB	*F.* × *ananassa* 'Jumbo × Bali'	Indonesia (local hybrid selection)	PV804957
SC	*F.* × *ananassa* 'Sachinoka'	Japan	PV804958
SH	*F.* × *ananassa* 'Sagahonoka'	Japan (Saga Prefecture)	PV804959
SL	*F.* × *ananassa* 'Seolhyang'	South Korea; 'Akihime' × 'Redpearl' (Hong & Eum, 2020)	PV804960
SG	*F.* × *ananassa* 'Saga'	Not documented	PV804961
LS	*F.* × *ananassa* 'Langensari'	Indonesia (local cultivar)	PV804962
DL	*F.* × *ananassa* 'Delife'	Not documented	PV804963
TO	*F.* × *ananassa* 'Tochiotome'	Japan (Tochigi Prefecture)	PV804964
FS	*F.* × *ananassa* 'Festival'	USA; 'Rosa Linda' × 'Oso Grande' (Chandler et al., 2000)	PV804965

Table 1 (*Continued*)

Code	Cultivar or Species	Source of Collection	GenBank Accession
MS	*F.* × *ananassa* 'Mayeonseol'	Not documented	PV804967
-	*F. iinumae*	Japan	AB073685
-	*F. vesca*	Portugal	HM850686
-	*F. vesca*	Italy	HQ619833
-	*F. vesca* subsp. *americana*	Canada	KJ840918
-	*F. vesca*	Canada	MG221050
-	*F. virginiana*	USA	KJ592952
-	*F. virginiana*	Canada	KJ840920
-	*F. virginiana*	Canada	MG221052
-	*F. chiloensis*	Canada	KX677569
-	*F. orientalis*	China	MN492802
-	*Potentilla indica* (outgroup)	Italy	HE966972

Note. Origin: country where the cultivar was bred. All 15 cultivars were collected in Indonesia (see Materials and Methods). Not documented: name used by local growers; breeding origin could not be traced. For GenBank accessions, origin is the country of the source plant

Amplification and Sequencing of the *matK* Gene

The *matK* region was amplified with the MyTaq Plant-PCR Kit (Bioline), *matK*3F (5'-TCC CTC TTC TTT GCA TTT ATT ACG-3'), and *matK*4R (5'-GCG TTA CAA AAT TTC ACT TTA GCC-3') primers from Potter et al. (2002). Each 30 µL reaction contained 12.5 µL of MyTaq Plant-PCR mix, 2 µL of each primer (10 µM), 5 µL of template DNA (10 ng/µL; 50 ng per reaction) and 8.5 µL of ddH_2O. Amplification was carried out in a T100 Thermal Cycler (Bio-Rad) with an initial denaturation at 95 °C for 5 min, followed by 35 cycles of denaturation at 95 °C for 30 s, annealing at 54.1 °C for 30 s and extension at 72 °C for 45 s, and a final extension at 72 °C for 5 min. The annealing temperature was selected after optimization of the PCR conditions. The amplification products were checked by 1% agarose gel electrophoresis at 100 V for 30 minutes using a 100 bp DNA ladder (Figure 2). Sequencing was performed in both directions with the same primers using an Applied Biosystems 3500 Genetic Analyzer at the Integrated Research and Testing Laboratory (LPPT), Universitas Gadjah Mada.

Sequence Preparation and Alignment

Forward and reverse reads were merged into a consensus using Geneious Prime to reduce read errors. Ambiguities were manually checked based on the peaks of the sequence graph. Consensus confirmation was performed via online BLAST to ensure species similarity. Sequence alignment was performed with MAFFT v7.505 (Katoh & Standley, 2013).

Because base calls are least reliable at read ends and the sequences differed in length, the alignment was trimmed to the region shared by all 15 cultivar sequences, and a further 20 bp were removed from each end. The final alignment was 555 bp long (positions 540-1094 of the *F. iinumae matK* sequence AB073685); two homopolymer length differences in *F. iinumae* were aligned as indels. The alignment was exported in FASTA, Nexus, and Phylip formats for further analysis.

Detection of Genetic Variation and Measurement of Genetic Distance

The number of polymorphic sites and haplotypes, haplotype diversity (Hd), and nucleotide diversity (π) (Nei, 1987) were calculated from the trimmed alignment using DnaSP 6 (Rozas et al., 2017). Genetic distance measurements were performed using the MEGA11 program (Tamura et al., 2021) and the Kimura 2-parameter model (Kimura, 1980) as a parameter method, since it calculates differences in transition and transversion changes (Nei & Kumar, 2000).

Haplotype Network and Phylogenetic Construction

Haplotype network construction followed the median-joining method (Bandelt et al., 1999), with each indel counted as one step; the network was drawn in Python (Matplotlib). The phylogenetic reconstruction process starts with selecting the best base substitution model using ModelFinder (Kalyaanamoorthy et al., 2017) under the Bayesian Information Criterion. The Maximum Likelihood (ML) method was applied using IQ-TREE v2.3.6 (Minh et al., 2020), where 1,000 ultrafast bootstrap replicates (Hoang et al., 2018) and 1,000 SH-aLRT replicates (Guindon et al., 2010) were used to assess the stability of the phylogenetic tree branches. Meanwhile, the Bayesian Inference (BI) method was carried out with MrBayes v3.2.7a (Ronquist et al., 2012) under the HKY model, using two runs of four chains for 2,000,000 generations, sampling every 1,000 generations, with a burn-in of 25% to eliminate unstable initial influences. Convergence was checked using the average standard deviation of split frequencies (< 0.01), the potential scale reduction factor (≈ 1.00) and effective sample sizes (> 200). *Potentilla indica* (HE966972) was used as the outgroup (Dobeš & Paule, 2010). The results of both methods were then visualized and edited, and nodes with SH-aLRT $\geq 80\%$, bootstrap $\geq 95\%$ and posterior probability ≥ 0.95 were regarded as well supported.

RESULTS AND DISCUSSIONS

DNA Quality

The DNA extraction in this study was subjected to qualitative and quantitative DNA tests. Qualitative tests showed that the isolated DNA had a high molecular weight, with some

samples exhibiting smears indicating partial degradation. DNA concentrations ranged from 15.2 to 82.0 ng/µL, and three samples ('Delife', 'Tochiotome', and 'Festival') exceeded 70 ng/µL. Only 'Mayeonseol' had an A260/A280 ratio ranging from 1.8-2.0; the other samples ranged from 1.34 to 1.79, which indicates protein or phenolic contamination (Table 2). Despite this, all 15 samples produced a single, clear amplicon (Figure 2) and were therefore retained for sequencing.

Table 2
DNA quality test results of 15 strawberry (F. × ananassa) cultivars

No.	Sample Code	Cultivar	Concentration (ng/µL)	Purity (A260/A280)
1.	CL	'Californica'	56.4	1.49
2.	KN	'Knia'	21.9	1.64
3.	MR	'Mencir'	25.6	1.63
4.	MH	'Moha'	21.8	1.63
5.	GL	'Geolhyang'	15.2	1.79
6.	JB	'Jumbo × Bali'	22.2	1.55
7.	SC	'Sachinoka'	21.2	1.69
8.	SH	'Sagahonoka'	21.7	1.56
9.	SL	'Seolhyang'	18.9	1.63
10.	SG	'Saga'	51.9	1.53
11.	LS	'Langensari'	39.3	1.73
12.	DL	'Delife'	73.5	1.34
13.	TO	'Tochiotome'	72.1	1.40
14.	FS	'Festival'	82.0	1.60
15.	MS	'Mayeonseol'	33.0	1.96

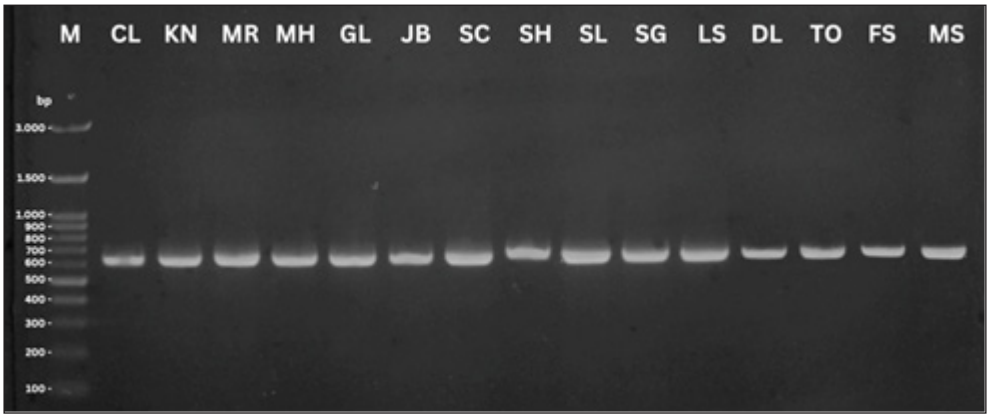


Figure 2. Electrophoregram of *matK* amplification products from 15 strawberry (*F. × ananassa*) cultivar samples on a 1% agarose gel. Note: M: Marker 100 bp DNA ladder, CL: 'Californica', KN: 'Knia', MR: 'Mencir', MH: 'Moha', GL: 'Geolhyang', JB: 'Jumbo × Bali', SC: 'Sachinoka', SH: 'Sagahonoka', SL: 'Seolhyang', SG: 'Saga', LS: 'Langensari', DL: 'Delife', TO: 'Tochiotome', FS: 'Festival', and MS: 'Mayeonseol'

Amplification and Sequencing of *matK* Gene

All 15 samples produced a single band of about 600-700 bp (Figure 2), in agreement with the expected product of about 630 bp between the primer sites. The consensus sequences were 599-627 bp long and were deposited in GenBank under accession numbers PV804952-PV804965 and PV804967 (Table 1). The sequences started and ended at different positions, and the outermost bases of several reads had low-quality or ambiguous calls. For example, 'Festival' carried ambiguity codes (K and W) near both ends, and 'Seolhyang' differed from the other cultivars by a single T/A call at the 16th base of its read. These end regions and the primer sites were removed before analysis, leaving a 555-bp alignment without missing data in any of the 26 sequences.

Genetic Variation and Genetic Distance

Within the 555-bp alignment, the 15 cultivar sequences were identical. No polymorphic site, indel or ambiguous position was found, and all cultivars belonged to a single haplotype (Hd = 0; $\pi = 0$). When the ten reference sequences were added, four variable sites and two indels (1 bp and 2 bp) were found (Table 3). All four variable sites were singletons, so none was parsimony-informative. Two substitutions and both indels occurred only in *F. iinumae* (AB073685), and two substitutions occurred only in *F. vesca* (HM850686). Transversions (three) outnumbered transitions (one). The 25 ingroup sequences formed three haplotypes (Hd = 0.157; $\pi = 0.00058$; Table 4).

K2P distances among the 15 cultivars were zero (Table 5). The cultivar haplotype was also identical to the *matK* sequences of *F. virginiana*, *F. chiloensis*, *F. orientalis* and three *F. vesca* accessions. Non-zero distances were found only between this haplotype and *F. iinumae* (0.0036) or *F. vesca* HM850686 (0.0036), and between these two accessions (0.0073). Distances to the outgroup *P. indica* ranged from 0.0548 to 0.0568.

Table 3
Variable sites in the 555-bp matK alignment of 25 Fragaria sequences, shown by haplotype

Haplotype	Position in Alignment	39	49	50	195	298	515	545
	Position in AB073685	**578**	**588**	**589**	**734**	**837**	**1054**	**1084**
H1	15 cultivars + 8 references	-	-	-	T	G	T	T
H2	*F. iinumae* (AB073685)	A	T	T	C	T	.	.
H3	*F. vesca* (HM850686)	-	-	-	.	.	G	G

Note. Positions are given along the 555-bp alignment and relative to the complete *matK* coding sequence of *F. iinumae* (AB073685); alignment position 1 corresponds to position 540 of AB073685. "." indicates the same base as H1 and "-" indicates a gap

Table 4
Polymorphism and diversity of the matK gene (555 bp) in strawberry cultivars grown in Indonesia and in all ingroup sequences

Data set	n	S	Pi	Ts	Tv	Indels	h	Hd	π
Cultivars grown in Indonesia	15	0	0	0	0	0	1	0.000	0.00000
All ingroup sequences	25	4	0	1	3	2 (3 sites)	3	0.157	0.00058

Note. n, number of sequences; S, variable sites; Pi, parsimony-informative sites; Ts, transitions; Tv, transversions; h, number of haplotypes; Hd, haplotype diversity; π, nucleotide diversity

Table 5
Kimura 2-parameter distances among matK haplotypes and the outgroup (555 bp)

	H1	H2	H3	*P. indica*
H1	-			
H2 (*F. iinumae*)	0.0036	-		
H3 (*F. vesca* HM850686)	0.0036	0.0073	-	
P. indica (outgroup)	0.0548	0.0568	0.0548	-

The absence of *matK* variation among the cultivars contrasts with earlier RAPD and CAPS studies, which detected polymorphism among strawberry cultivars in Magelang (Arif et al., 2019; Aristya et al., 2019). This difference is expected. RAPD and CAPS markers survey many nuclear loci, whereas *matK* is a single, slowly evolving plastid gene inherited through the maternal line. Low plastid variation is typical of *Fragaria* (Cheng et al., 2017; Li et al., 2021; Njuguna et al., 2013), and the whole octoploid clade descends from one maternal lineage (Njuguna et al., 2013; Song et al., 2023). The genetic distances obtained here therefore describe maternal lineage and cannot be used as a measure of differences between cultivars.

Haplotype Network based on the *matK* Gene

The 25 ingroup sequences formed three haplotypes (Table 6; Figure 3). H1 was shared by all 15 cultivars and by eight reference accessions of *F. virginiana*, *F. chiloensis*, *F. vesca* and *F. orientalis*. H2 (*F. iinumae*, AB073685) was separated from H1 by two substitutions and two indels, and H3 (*F. vesca*, HM850686) by two substitutions. The median-joining method is effective because it simplifies the network, generating a more refined representation of evolutionary relationships (Leigh & Bryant, 2015). The circle size on each haplotype indicates the number of samples represented, with H1 having the largest size, and the lines and short perpendicular marks indicate mutational steps between haplotypes.

Table 6
Haplotypes of 25 strawberry (Fragaria spp.) *sequences based on the matK gene*

Haplotype	**Number of Samples**	**Sample Code/ *Accession***
H1	23	CL, KN, MR, MH, GL, JB, SC, SH, SL, SG, LS, DL, TO, FS, MS; KJ592952, KJ840920, MG221052 (*F. virginiana*); KX677569 (*F. chiloensis*); MG221050, HQ619833, KJ840918 (*F. vesca*); MN492802 (*F. orientalis*)
H2	1	AB073685 (*F. iinumae*)
H3	1	HM850686 (*F. vesca*)

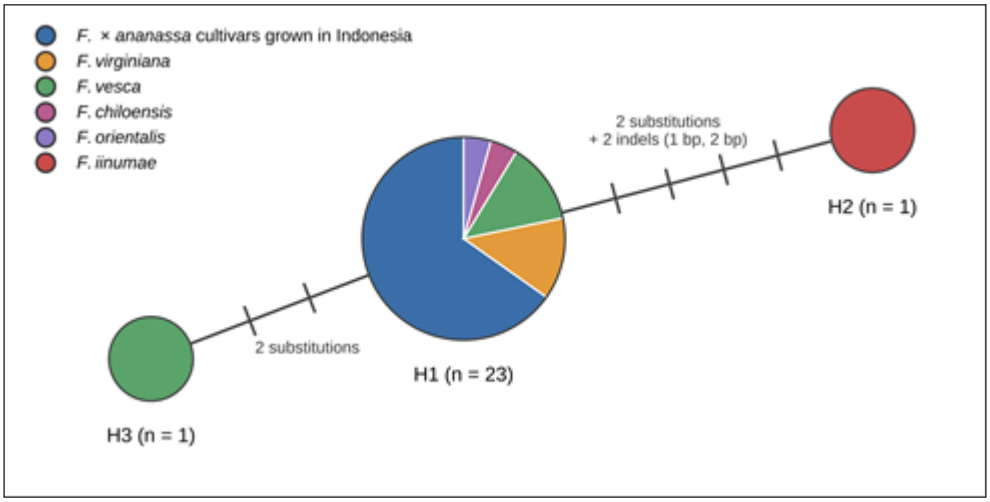


Figure 3. Median-joining haplotype network of 25 *Fragaria* sequences based on the *matK* gene (555 bp). Circle size is proportional to the number of sequences, colours indicate taxa, and hatch marks indicate mutational steps (substitutions or indels)

Because the cultivars share H1 with accessions of several species collected in different countries, the network cannot show the geographic origin of the cultivars; the country of a GenBank accession only describes where that reference plant was collected. The shared haplotype is better explained by the origin of the crop. Cultivated strawberry derives from hybrids between *F. virginiana* and *F. chiloensis* (Hardigan et al., 2021; Liston et al., 2014), and the plastid sequences of these species and of *F.* × *ananassa* are very similar (Cheng et al., 2017; Honjo et al., 2009). According to Hanif and Ashari (2012), strawberries cultivated in Indonesia are hybrid strawberries derived from various types of strawberries. The cultivars examined here include Japanese, Korean and American releases and local selections (Table 1), yet all share H1, so their pedigrees, for example 'Seolhyang' ('Akihime' × 'Redpearl'; Hong & Eum, 2020) and 'Festival' ('Rosa Linda' × 'Oso Grande'; Chandler et al., 2000), cannot be recovered from *matK*.

Strawberry Phylogenetic Analysis Based on the *matK* Gene

Phylogenetic tree building was performed with one outgroup, *Potentilla indica* (HE966972), a close relative of the genus *Fragaria* in the Rosaceae family (Dobeš & Paule, 2010). The best base substitution model was HKY+F (BIC = 1851.80). The HKY85 model discriminates between transition and transversion rates using the κ (kappa) parameter and allows for unequal base frequencies (Hasegawa et al., 1985). The ML and BI methods produced the same topology (Figure 4). *Fragaria vesca* HM850686 diverged first, and the remaining 24 sequences (H1 and *F. iinumae*) formed one group with weak support (SH-aLRT 77.5%, bootstrap 62%, posterior probability 0.84). All other branches were of zero length, so the 15 cultivars and the H1 references formed a single unresolved group (polytomy). The BI runs converged well (ASDSF = 0.0005; PSRF = 1.000-1.001; minimum ESS = 1,202). Using the ML method, the bootstrap value on the branching of the phylogenetic tree can provide information on the confidence level in the proposed kinship relationship, and posterior probability values range from 0-1, with values tending toward one indicating high confidence (Ronquist et al., 2012). The low values obtained here show that the grouping is not well supported.

The polytomy is a direct result of identical sequences and shows that *matK* carries no phylogenetic signal at the cultivar level. The grouping of *F. iinumae* with H1 should not be interpreted as a close relationship. *Fragaria iinumae* is a diploid species from Japan that represents one of the oldest lineages of the genus and one of the nuclear progenitors of the octoploids (Edger et al., 2019; Njuguna et al., 2013; Rousseau-Gueutin et al., 2009). Its position here rests on a few unique changes and weak support, and it reflects the limited information in a 555-bp alignment rather than true ancestry. Whole-plastome studies place the maternal origin of the octoploid clade, including the cultivated strawberry, in western North American *F. vesca* subsp. *bracteata* (Njuguna et al., 2013; Song et al., 2023). The *F. vesca* accession that differed from H1 probably reflects plastid variation within this widespread diploid species (Liston et al., 2014).

These data have practical value beyond the cultivars studied. First, *matK* is not suitable for authenticating strawberry cultivars or inferring their origin, although it can still confirm that a sample belongs to *Fragaria*. Second, the apparent polymorphisms at unequal read ends found in this study show that Sanger barcode data need simple quality checks before interpretation: inspecting chromatograms, trimming read ends to the region shared by all sequences, reporting the effective alignment length, and giving diversity indices such as Hd and π. Third, the marker has to match the research question (Table 7). Variable plastid regions such as *trnL-trnF*, *trnR-rrn5* or the intergenic loci identified by Li et al. (2021) are better suited to maternal lineages and species, whereas cultivar identification requires nuclear markers such as SSRs (Honjo et al., 2011), CAPS (Arif et al., 2019), DArT (Sánchez-Sevilla et al., 2015) or genome-wide SNPs (Hardigan et al., 2021).

Characterizing germplasm with such markers can also reveal genetic diversity useful for breeding (Dávalos-González et al., 2023).

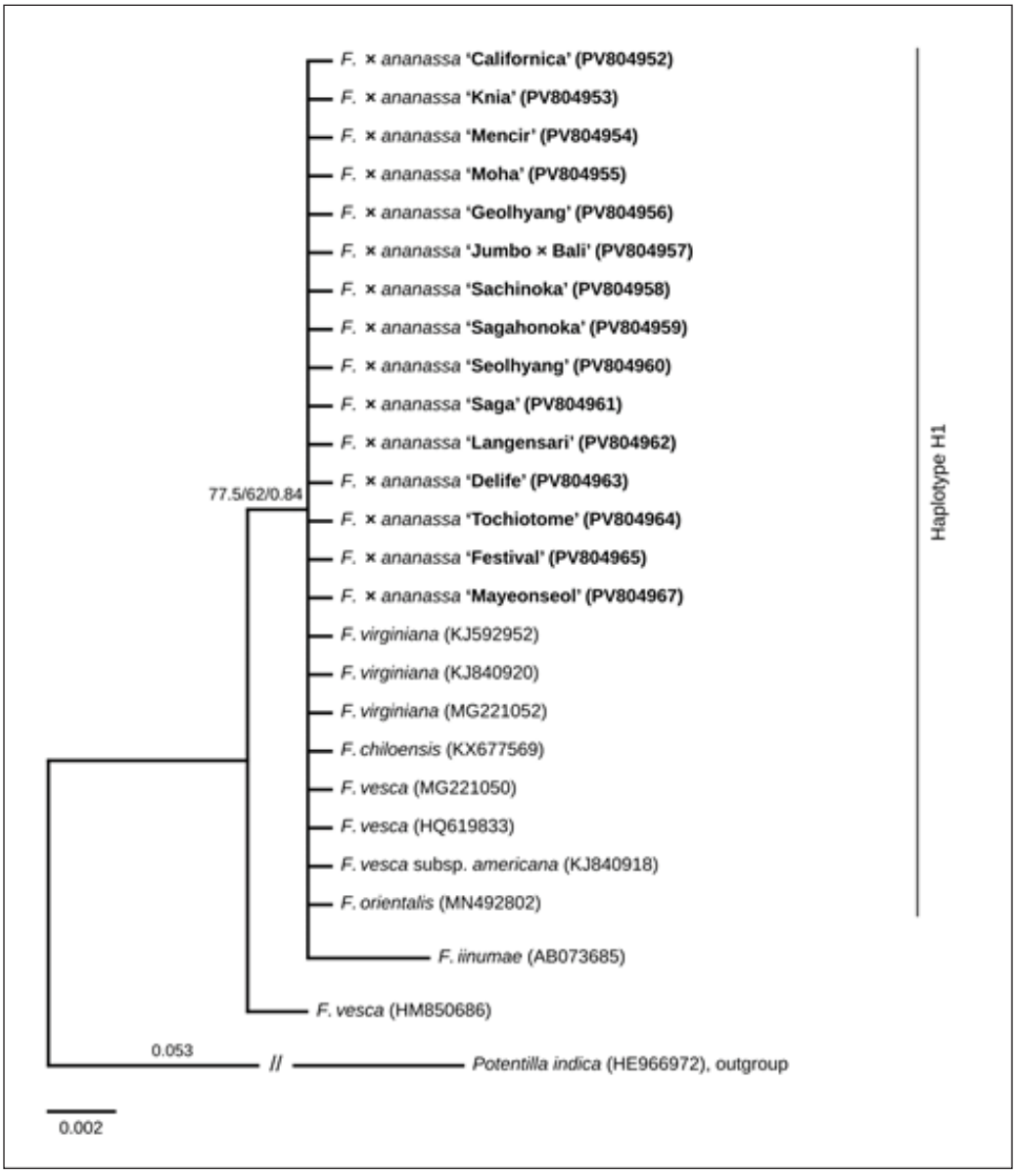


Figure 4. Phylogenetic tree of 25 *Fragaria matK* sequences (555 bp) and the outgroup *Potentilla indica* reconstructed with Maximum Likelihood (IQ-TREE, HKY+F) and Bayesian Inference (MrBayes, HKY); both methods gave the same topology. The node label shows SH-aLRT (%)/ultrafast bootstrap (%)/posterior probability; the remaining branches had zero length. Cultivars grown in Indonesia are shown in bold. The outgroup branch is shortened (//); the scale bar shows substitutions per site

Table 7
Suggested molecular markers for different purposes in strawberry research, based on this study and the literature

Purpose	Suitable Markers	Remarks
Confirming genus or species identity	*matK*, *rbcL* (core plant barcodes)	Resolve the genus; little or no resolution among octoploid taxa and cultivars (CBOL Plant Working Group, 2009; this study)
Tracing maternal lineage	*trnL-trnF*, *trnR-rrn5*; whole plastome	Separate *F. virginiana*- and *F. chiloensis*-type maternal lineages (Honjo et al., 2009; Njuguna et al., 2013)
Separating wild *Fragaria* species	*trnS-trnG*, *trnR-atpA*, *trnC-petN*, *rbcL-accD*, *psbE-petL*	Most variable plastid loci among wild species (Li et al., 2021)
Cultivar identification and fingerprinting	SSR, CAPS	Low-cost and reproducible; used for Japanese and Indonesian cultivars (Arif et al., 2019; Honjo et al., 2011)
Diversity, population structure and breeding	DArT, SNP arrays, resequencing	Genome-wide resolution of cultivar diversity (Fan & Whitaker, 2024; Hardigan et al., 2021; Sánchez-Sevilla et al., 2015)

CONCLUSION

Genetic variation analyzed from the *matK* gene showed no variation among the 15 strawberry cultivars grown in Indonesia, and the genetic distance among cultivars was zero. In addition, the distribution pattern obtained from the haplotype network showed that all cultivars share one haplotype with *F. virginiana*, *F. chiloensis*, *F. vesca* and *F. orientalis*, which shows a common maternal lineage rather than geographic origin. Finally, the phylogenetic relationship analyzed by *matK* gene sequences using the ML and BI methods showed that the cultivars form one unresolved group. Therefore, *matK* cannot distinguish these cultivars.

IMPLICATION OF THE STUDY

As the first *matK* dataset for strawberry cultivars grown in Indonesia, these sequences provide reference data and a clear basis for choosing more informative markers. The results also show that *matK* can confirm that a sample belongs to *Fragaria*, but plastid barcodes are not suitable for authenticating strawberry cultivars or inferring their origin.

LIMITATION AND RECOMMENDATION FOR FUTURE RESEARCH

This study was limited to one plastid locus and few plants per cultivar. Nuclear markers such as SSRs or SNPs, together with documented pedigrees, are recommended for cultivar identification and breeding programs.

DATA AVAILABILITY

The *matK* sequences generated in this study are available in GenBank under accession numbers PV804952-PV804965 and PV804967.

ACKNOWLEDGEMENT

The authors would like to thank the Government of the Republic of Indonesia, specifically the Ministry of Higher Education, Research, and Technology, for funding Fundamental Research Grant Number 2465/UN1/DITLIT/DitLit/PT.01.03/2025. Also, the authors would like to thank everyone involved in preparing this article.

REFERENCES

Arif, M. F., Aristya, G. R., & Kasiamdari, R. S. (2019). Genetic diversity of strawberry cultivars in Banyuroto, Magelang, Indonesia based on cleaved amplified polymorphic sequence. *Biodiversitas Journal of Biological Diversity, 20*(6), 1721-1728. https://doi.org/10.13057/biodiv/d200631

Aristya, G. R., Kasiamdari, R. S., Setyoningrum, R., & Larasati, B. (2019). Genetic variations of strawberry cultivars of *Fragaria* × *ananassa* and *Fragaria vesca* based on RAPD. *Biodiversitas Journal of Biological Diversity, 20*(3), 770-775. https://doi.org/10.13057/biodiv/d200322

Azizah, U. D. L., Yulianti, F., Adiredjo, A. L., & Sitawati. (2019). Genetic relationship analysis of strawberry (*Fragaria* sp.) germplasm based on morphology character and random amplified polymorphic DNA (RAPD). *Plantropica: Journal of Agricultural Science, 4*(1), 77-85. https://doi.org/10.21776/ub.jpt.2019.004.1.9

Bandelt, H. J., Forster, P., & Röhl, A. (1999). Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution, 16*(1), 37-48. https://doi.org/10.1093/oxfordjournals.molbev.a026036

CBOL Plant Working Group. (2009). A DNA barcode for land plants. *Proceedings of the National Academy of Sciences of the United States of America, 106*(31), 12794-12797. https://doi.org/10.1073/pnas.0905845106

Chandler, C. K., Legard, D. E., Dunigan, D. D., Crocker, T. E., & Sims, C. A. (2000). 'Strawberry Festival' strawberry. *HortScience, 35*(7), 1366-1367. https://doi.org/10.21273/HORTSCI.35.7.1366

Cheng, H., Li, J., Zhang, H., Cai, B., Gao, Z., Qiao, Y., & Mi, L. (2017). The complete chloroplast genome sequence of strawberry (*Fragaria* × *ananassa* Duch.) and comparison with related species of Rosaceae. *PeerJ, 5*, Article e3919. https://doi.org/10.7717/peerj.3919

Dávalos-González, P. A., Aguilar-García, R., Rodríguez-Guillén, A., & Jofre-y-Garfias, A. E. (2023). The genetic diversity of strawberry species, the underutilized gene pool and the need for cultivars with new quality and agronomic attributes. In N. E. Yaşa Kafkas (Ed.), *Recent studies on strawberries*. IntechOpen. https://doi.org/10.5772/intechopen.102962

Dharmayanti, N. L. P. I. (2011). Filogenetika molekuler: Metode taksonomi organisme berdasarkan sejarah evolusi [Molecular phylogenetics: A method of organismal taxonomy based on evolutionary history]. *Wartazoa, 21*(1), 1-10.

Direktorat Jenderal Hortikultura. (2025). *Angka tetap hortikultura tahun 2024* [Fixed horticultural statistics 2024].

Dobeš, C., & Paule, J. (2010). A comprehensive chloroplast DNA-based phylogeny of the genus *Potentilla* (Rosaceae): Implications for its geographic origin, phylogeography and generic circumscription. *Molecular Phylogenetics and Evolution, 56*(1), 156-175. https://doi.org/10.1016/j.ympev.2010.03.005

Edger, P. P., Poorten, T. J., VanBuren, R., Hardigan, M. A., Colle, M., McKain, M. R., Smith, R. D., Teresi, S. J., Nelson, A. D. L., Wai, C. M., Alger, E. I., Bird, K. A., Yocca, A. E., Pumplin, N., Ou, S., Ben-Zvi, G., Brodt, A., Baruch, K., Swale, T., . . . Knapp, S. J. (2019). Origin and evolution of the octoploid strawberry genome. *Nature Genetics, 51*(3), 541-547. https://doi.org/10.1038/s41588-019-0356-4

Fan, Z., & Whitaker, V. M. (2024). Genomic signatures of strawberry domestication and diversification. *The Plant Cell, 36*(5), 1622-1636. https://doi.org/10.1093/plcell/koad314

Farooq, Q., Shakir, M., Ejaz, F., Zafar, T., Durrani, K., & Ullah, A. (2020). Role of DNA barcoding in plant biodiversity conservation. *Scholars International Journal of Biochemistry, 3*(3), 48-52. https://doi.org/10.36348/sijb.2020.v03i03.002

Guindon, S., Dufayard, J. F., Lefort, V., Anisimova, M., Hordijk, W., & Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. *Systematic Biology, 59*(3), 307-321. https://doi.org/10.1093/sysbio/syq010

Hanif, Z., & Ashari, H. (2012). Sebaran stroberi (*Fragaria* × *ananassa*) di Indonesia [Distribution of strawberry (*Fragaria* × *ananassa*) in Indonesia]. In *Prosiding Seminar Nasional Pekan Inovasi Teknologi Hortikultura Nasional* (pp. 87-95). Pusat Penelitian dan Pengembangan Hortikultura. https://doi.org/10.13140/RG.2.1.2110.6089

Hardigan, M. A., Lorant, A., Pincot, D. D. A., Feldmann, M. J., Famula, R. A., Acharya, C. B., Lee, S., Verma, S., Whitaker, V. M., Bassil, N., Zurn, J., Cole, G. S., Bird, K., Edger, P. P., & Knapp, S. J. (2021). Unraveling the complex hybrid ancestry and domestication history of cultivated strawberry. *Molecular Biology and Evolution, 38*(6), 2285-2305. https://doi.org/10.1093/molbev/msab024

Hasegawa, M., Kishino, H., & Yano, T. (1985). Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution, 22*(2), 160-174. https://doi.org/10.1007/BF02101694

Hoang, D. T., Chernomor, O., von Haeseler, A., Minh, B. Q., & Vinh, L. S. (2018). UFBoot2: Improving the ultrafast bootstrap approximation. *Molecular Biology and Evolution, 35*(2), 518-522. https://doi.org/10.1093/molbev/msx281

Hong, S. J., & Eum, H. L. (2020). Determination of the harvest date and ripening phase of 'Seolhyang' strawberry. *Protected Horticulture and Plant Factory, 29*(1), 62-72. https://doi.org/10.12791/KSBEC.2020.29.1.62

Honjo, M., Kataoka, S., Yui, S., Morishita, M., Kunihisa, M., Yano, T., Hamano, M., & Yamazaki, H. (2009). Maternal lineages of the cultivated strawberry, *Fragaria* × *ananassa*, revealed by chloroplast DNA variation. *HortScience, 44*(6), 1562-1565. https://doi.org/10.21273/HORTSCI.44.6.1562

Honjo, M., Nunome, T., Kataoka, S., Yano, T., Yamazaki, H., Hamano, M., Yui, S., & Morishita, M. (2011). Strawberry cultivar identification based on hypervariable SSR markers. *Breeding Science, 61*(4), 420-425. https://doi.org/10.1270/jsbbs.61.420

Joly, S., Bryant, D., & Lockhart, P. J. (2015). Flexible methods for estimating genetic distances from single nucleotide polymorphisms. *Methods in Ecology and Evolution, 6*(8), 938-948. https://doi.org/10.1111/2041-210X.12343

Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A., & Jermiin, L. S. (2017). ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nature Methods, 14*(6), 587-589. https://doi.org/10.1038/nmeth.4285

Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Molecular Biology and Evolution, 30*(4), 772-780. https://doi.org/10.1093/molbev/mst010

Kimura, M. (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution, 16*(2), 111-120. https://doi.org/10.1007/BF01731581

Koryati, T., Ningsih, H., Erdiandini, I., Paulina, M., Firgiyanto, R., Junairiah, & Sari, V. K. (2022). *Pemuliaan tanaman* [Plant breeding]. Yayasan Kita Menulis.

Leigh, J. W., & Bryant, D. (2015). POPART: Full-feature software for haplotype network construction. *Methods in Ecology and Evolution, 6*(9), 1110-1116. https://doi.org/10.1111/2041-210X.12410

Li, C., Cai, C., Tao, Y., Sun, Z., Jiang, M., Chen, L., & Li, J. (2021). Variation and evolution of the whole chloroplast genomes of *Fragaria* spp. (Rosaceae). *Frontiers in Plant Science, 12*, Article 754209. https://doi.org/10.3389/fpls.2021.754209

Liston, A., Cronn, R., & Ashman, T. L. (2014). *Fragaria*: A genus with deep historical roots and ripe for evolutionary and ecological insights. *American Journal of Botany, 101*(10), 1686-1699. https://doi.org/10.3732/ajb.1400140

Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., von Haeseler, A., & Lanfear, R. (2020). IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. *Molecular Biology and Evolution, 37*(5), 1530-1534. https://doi.org/10.1093/molbev/msaa015

Nei, M. (1987). *Molecular evolutionary genetics*. Columbia University Press. https://doi.org/10.7312/nei-92038

Nei, M., & Kumar, S. (2000). *Molecular evolution and phylogenetics*. Oxford University Press. https://doi.org/10.1093/oso/9780195135848.001.0001

Njuguna, W., Liston, A., Cronn, R., Ashman, T. L., & Bassil, N. (2013). Insights into phylogeny, sex function and age of *Fragaria* based on whole chloroplast genome sequencing. *Molecular Phylogenetics and Evolution, 66*(1), 17-29. https://doi.org/10.1016/j.ympev.2012.08.026

Potter, D., Gao, F., Bortiri, P. E., Oh, S. H., & Baggett, S. (2002). Phylogenetic relationships in Rosaceae inferred from chloroplast *matK* and *trnL-trnF* nucleotide sequence data. *Plant Systematics and Evolution, 231*(1-4), 77-89. https://doi.org/10.1007/s006060200012

Qarni, A., Muhammad, K., Wahab, A., Ali, A., Khizar, C., Ullah, I., Kazmi, A., Sultana, T., Hameed, A., Younas, M., & Rahimi, M. (2022). Molecular characterization of wild and cultivated strawberry (*Fragaria* × *ananassa*) through DNA barcode markers. *Genetics Research, 2022*, Article 9249561. https://doi.org/10.1155/2022/9249561

Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D. L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M. A., & Huelsenbeck, J. P. (2012). MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology, 61*(3), 539-542. https://doi.org/10.1093/sysbio/sys029

Rousseau-Gueutin, M., Gaston, A., Aïnouche, A., Aïnouche, M. L., Olbricht, K., Staudt, G., Richard, L., & Denoyes-Rothan, B. (2009). Tracking the evolutionary history of polyploidy in *Fragaria* L. (strawberry): New insights from phylogenetic analyses of low-copy nuclear genes. *Molecular Phylogenetics and Evolution, 51*(3), 515-530. https://doi.org/10.1016/j.ympev.2008.12.024

Rozas, J., Ferrer-Mata, A., Sánchez-DelBarrio, J. C., Guirao-Rico, S., Librado, P., Ramos-Onsins, S. E., & Sánchez-Gracia, A. (2017). DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Molecular Biology and Evolution, 34*(12), 3299-3302. https://doi.org/10.1093/molbev/msx248

Sambrook, J., & Russell, D. W. (2001). *Molecular cloning: A laboratory manual* (3rd ed.). Cold Spring Harbor Laboratory Press.

San Mauro, D., & Agorreta, A. (2010). Molecular systematics: A synthesis of the common methods and the state of knowledge. *Cellular & Molecular Biology Letters, 15*(2), 311-341. https://doi.org/10.2478/s11658-010-0010-8

Sánchez-Sevilla, J. F., Horvath, A., Botella, M. A., Gaston, A., Folta, K., Kilian, A., Denoyes, B., & Amaya, I. (2015). Diversity Arrays Technology (DArT) marker platforms for diversity analysis and linkage mapping in a complex crop, the octoploid cultivated strawberry (*Fragaria* × *ananassa*). *PLOS ONE, 10*(12), Article e0144960. https://doi.org/10.1371/journal.pone.0144960

Sapei, L., & Hwa, L. (2014). Study on the kinetics of vitamin C degradation in fresh strawberry juices. *Procedia Chemistry, 9*, 62-68. https://doi.org/10.1016/j.proche.2014.05.008

Savolainen, V., Cowan, R. S., Vogler, A. P., Roderick, G. K., & Lane, R. (2005). Towards writing the encyclopaedia of life: An introduction to DNA barcoding. *Philosophical Transactions of the Royal Society B: Biological Sciences, 360*(1462), 1805-1811. https://doi.org/10.1098/rstb.2005.1730

Siagian, H., Hasibuan, S., & Suswati, S. (2016). Aplikasi benzyl amino purin (BAP) terhadap pertumbuhan dan produksi stroberi (*Fragaria* × *ananassa* var. Duchesne) dari sumber bibit yang berbeda [Application of benzyl amino purine (BAP) on the growth and production of strawberry from different seedling sources]. *Agrotekma: Jurnal Agroteknologi dan Ilmu Pertanian, 1*(1). https://doi.org/10.31289/agr.v1i1.1106

Song, Y., Li, C., Liu, L., Hu, P., Li, G., Zhao, X., & Zhou, H. (2023). The population genomic analyses of chloroplast genomes shed new insights on the complicated ploidy and evolutionary history in *Fragaria*. *Frontiers in Plant Science, 13*, Article 1065218. https://doi.org/10.3389/fpls.2022.1065218

Sun, J., Sun, R., Liu, H., Chang, L., Li, S., Zhao, M., Shennan, C., Lei, J., Dong, J., Zhong, C., Xue, L., Gao, Y., Wang, G., & Zhang, Y. (2021). Complete chloroplast genome sequencing of ten wild *Fragaria* species in China provides evidence for phylogenetic evolution of *Fragaria*. *Genomics, 113*(3), 1170-1179. https://doi.org/10.1016/j.ygeno.2021.01.027

Tamura, K., Stecher, G., & Kumar, S. (2021). MEGA11: Molecular Evolutionary Genetics Analysis version 11. *Molecular Biology and Evolution, 38*(7), 3022-3027. https://doi.org/10.1093/molbev/msab120

Udensi, O. U., Ita, E. E., Ikpeme, E. V., Ubi, G., & Emeagi, L. I. (2017). Sequence analysis of maturase K (*matK*): A chloroplast-encoding gene in some selected pulses. *Global Journal of Pure and Applied Sciences, 23*(2), 213-230. https://doi.org/10.4314/gjpas.v23i2.2

Vondracek, K., Altpeter, F., Liu, T., & Lee, S. (2024). Advances in genomics and genome editing for improving strawberry (*Fragaria* × *ananassa*). *Frontiers in Genetics, 15*, Article 1382445. https://doi.org/10.3389/fgene.2024.1382445